AFRL-AFOSR-CL-TR-2017-0004

Fluid analysis of network content dissemination and cloud systems

Fernando Paganini
UNIVERSIDAD ORT URUGUAY

03/06/2017
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ IOS
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 10-03-2017	**2. REPORT TYPE** Final	**3. DATES COVERED** *(From - To)* 01 Sep 2015 to 30 Nov 2016

4. TITLE AND SUBTITLE
Fluid analysis of network content dissemination and cloud systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-15-1-0183

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Fernando Paganini

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSIDAD ORT URUGUAY
CUAREIM 1451
MONTEVIDEO, 11100 UY

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/SOARD
U.S. Embassy Santiago
Av. Andres Bello 2800
Santiago, Chile

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR IOS

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-CL-TR-2017-0004

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The project was concerned with optimizing performance in complex network content dissemination and cloud systems. We employed tools of queueing theory, convex optimization and control theory, to study how to disseminate content in a distributed network, managing tradeoffs in efficiency, energy, and resource allocation. Fluid models provide a tractable path to represent these high dimensional problems, retaining accuracy in key performance questions. A first line of work, initiated in our previous AFOSR/SOARD project, concerns peer-to-peer dissemination in wireless ad-hoc networks. We focus on the necessary tradeoff between an efficient use of the network substrate, and the necessary reciprocity between peers, aspects that may be in conflict in the wireless setting. Our results published in [5] use convex optimization to formulate a relevant tradeoff, and propose decentralized algorithms which involve peer-to-peer interactions, and are shown to converge to the corresponding tradeoff point. A second line of contributions referred to the optimization of cache systems, a widespread method of content dissemination. In [2, 3] we address the question of which files to cache and its impact on performance; we worked in the setting of time-to-live (TTL) caching, where the decision involves a choice of timer for each stored content, and its relative popularity must be considered. We formulate a relevant optimization problem, and solve it in cases of practical interest. Numerous insights on practical caching mechanisms result from this mathematical analysis. Extensions of the method to networks of caches were tackled in [4]. A third direction concerned cloud computing and server systems, where processing resources may be adjusted dynamically in real time. The main question is how to control active service capacity, and how to allot it to current jobs, to pursue relevant performance.

15. SUBJECT TERMS
peer-to-peer networks, network cloud, "peer-to-peer networks" and "data centers", "performance analysis' and "network cloud", EOARD

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			MARTINEZ, MICHAEL
Unclassified	Unclassified	Unclassified	SAR	6	**19b. TELEPHONE NUMBER** *(Include area code)* 571-289-5167

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

FINAL PERFORMANCE REPORT: 1 Sep 2015 to 30 Nov 2016

AFOSR GRANT NUMBER: FA9550-15-1-0183

TITLE: Fluid analysis of network content dissemination and cloud systems

PI: Fernando Paganini
Universidad ORT,
Cuareim 1451, Montevideo, Uruguay

Abstract

The project was concerned with optimizing performance in complex network content dissemination and cloud systems. We employed tools of queueing theory, convex optimization and control theory, to study how to disseminate content in a distributed network, managing tradeoffs in efficiency, energy, and resource allocation. Fluid models provide a tractable path to represent these high dimensional problems, retaining accuracy in key performance questions.

A first line of work, initiated in our previous AFOSR/SOARD project, concerns peer-to-peer dissemination in wireless ad-hoc networks. We focus on the necessary tradeoff between an efficient use of the network substrate, and the necessary reciprocity between peers, aspects that may be in conflict in the wireless setting. Our results published in [5] use convex optimization to formulate a relevant tradeoff, and propose decentralized algorithms which involve peer-to-peer interactions, and are shown to converge to the corresponding tradeoff point.

A second line of contributions referred to the optimization of cache systems, a widespread method of content dissemination. In [2, 3] we address the question of which files to cache and its impact on performance; we worked in the setting of time-to-live (TTL) caching, where the decision involves a choice of timer for each stored content, and its relative popularity must be considered. We formulate a relevant optimization problem, and solve it in cases of practical interest. Numerous insights on practical caching mechanisms result from this mathematical analysis. Extensions of the method to networks of caches were tackled in [4].

A third direction concerned cloud computing and server systems, where processing resources may be adjusted dynamically in real time. The main question is how to control active service capacity, and how to allot it to current jobs, to pursue relevant performance objectives, in particular: fast response, energy savings, and predictable use of resources. An initial study which concerns smoothing of service capacity and has implications on power is [1]. More recently we have tackled the question of speed scaling in a cloud system, and shown in [6] that important performance gains can be obtained through fluid models and a control systems perspective.

Accomplishments/New Findings

We report separately on the areas of research indicated in the abstract.

1. Wireless P2P dissemination.

The efficient dissemination of a large content file among a set of network nodes requires the orchestration of multiple transfers, within the constraints of the communication substrate. In unstructured or aggressive environments where wireless ad-hoc networks are deployed, such dissemination must be arranged without a central network planner, and efficiency must be pursued by decentralized algorithms. Among them, the peer-to-peer (P2P) approach in which nodes exchange pieces of the file is an attractive option; however it has mostly been studied in a very different context, the wired Internet. In our work [5], continued from our previous project we investigated the challenges of P2P dissemination in the wireless substrate. In particular:

- Peering choices must take into account channel bandwidth. A tradeoff arises between the most efficient channel use and the requirement for peer reciprocity in the exchange. Our results formulate this mathematical tradeoff and decentralized algorithms to adjudicate it.

- Interference between wireless channels implies peer interactions are no longer independent, so a scheduling question arises. We show how to solve our efficiency/reciprocity tradeoff in this setting, by decoupling the problem through dual decomposition between the peer and medium access layers; the latter problem can be approximated by a Markov approximation scheme.

2. Optimizing cache systems.

On the opposite end of the dissemination question are content distribution networks (CDNs). These are highly structured networks of caches, maintained to hold the content of most interest closer to user demand, so as to reduce latency in the response time to download requests. A central question here is the decision of which content to store in these caches, and for how long, to obtain the best performance under distributed cache management. Our contributions to this problem were:

- In [2, 3] we investigated the optimization of storage time decisions in time-to-live (TTL) caches. In this arrangement, if a cache receives a request for a certain file, it keeps a copy and starts a timer, upon whose expiration the file is evicted, unless a new request is received first. The main decision variable is the selection of the timer value, which should depend on the popularity of the file. In our approach, for a renewal process of requests we express the occupation and "hit" probabilities as a function of the timer choice; the former is the fraction of time spent in the cache, the latter is the probability an arriving request will find the file in the cache. A natural optimization problem is formulated: maximize the hit probability for a given cache size. We show the solution depends on the request arrival process, in particular the *hazard rate* of the inter-arrival time distribution. In particular the alternative of caching permanently the most popular files need *not* be the optimal, it fails for the important case of bursty, heavy-tailed arrival process. Focusing on this case and on a Zipf-law for file popularities, we characterize the optimal policy and simplify its description by an appropriate fluid limit. This analysis, while applicable to the optimal TTL policy, also uncovers properties of other popular cache replacement policies, such as the LFU (least frequently used) and LRU (least recently used) rules for file eviction.

- In [4] we considered the situation of *arrays* of caches. In this case, rather than a purely autonomous operation for each cache, we consider the option of coordinating between them by moving files from one cache to another. This includes both *pulling* popular files gradually down the hierarchy, from the far-away repositories to the request sites, and also *pushing* evicted content upstream, in an attempt to keep it within a few hops of the requests. A number of alternatives in this regard have been identified. In our work we carried out experimental studies with the LRU policy, and analytical ones with TTL caches, characterizing parameter settings where the cache system achieves an efficient operation.

3. Speed scaling in server and cloud systems.

Current-day data centers and cloud computing systems are centralized facilities handling a large and aggregate load, varying in time. Rather than having them on at maximum capacity all the time, their processing speed can be scaled in a dynamic manner. This feature, together with scheduling decisions on arriving jobs, allow trading off response time with other desirable features, such as saving energy or producing a smoother usage profile.

The study of queueing systems under variable service capacity is comparatively under-developed; we have contributed to it in this project.

- In [1] we analyzed a system of deferrable jobs, whose *individual* service rate is controlled by a central entity. The main objective is to reduce the *variability* of the aggregate service capacity, subject to the constraint of meeting job deadlines. We analyze various control policies for this objective, with tools of queueing theory. The application to power systems is highlighted, where power variations are undesirable; but controlling variability has other applications as well.

- A different but related problem, known in the literature as “speed scaling” or “right-sizing”, is to dynamically adapt the *aggregate* service capacity to queue occupation, and then use some scheduling discipline to allocate this capacity amongst jobs present. Relevant concerns are energy consumption and per-job response time, and their robustness to variations in exogenous load. In our paper [6] we cast the problem in the setting of feedback control, using a fluid model of the queueing system; in this framework the problem is of designing a controller to track the exogenous demand, and the prior work can be seen as restricting the controller to a static function. By allowing for a dynamic controller, in particular a proportional-integral law, we show how the relevant performance tradeoff can be improved. We further indicate a discrete server implementation of this control law, based on a mix of dedicated servers and pooled helpers; its performance was evaluated analytically and by simulation.

Publications
(available from http://fi.ort.edu.uy/2243/17/publications.html)

[1] A. Ferragut, F. Paganini "Queueing analysis of service deferrals for load management in power systems", *Allerton Conference*, Monticello, IL, Oct 2015.

[2] A. Ferragut, I. Rodríguez F. Paganini, "Optimizing TTL Caches under Heavy-Tailed Demands", *ACM SIGMETRICS*, Antibes Juan-les-Pins, France, Jun 2016.

[3] A. Ferragut, I. Rodríguez F. Paganini, "Optimal TTL caching policies under general heavy tailed arrival processes", *Stochastic Networks Conference* , San Diego, CA, Jun 2016.

[4] I. Rodríguez, A. Ferragut, F. Paganini, "Improving performance of multiple-level cache systems", *ACM SIGCOMM-LANCOMM*, Florianopolis, Brazil, Aug 2016, pp. 37-39.

[5] F. Paganini, M. Zubeldía, A. Ferragut, "Reciprocity and Efficiency in Peer Exchange of Wireless Nodes through Convex Optimization", *IEEE Transactions on Network Science and Engineering*, Vol 3, No.4, pp. 257-270, Oct-Dec 2016.

[6] D. Goldsztajn, A. Ferragut, F. Paganini, "A feedback control approach to dynamic speed scaling in computing systems", *Conference on Information Sciences and Systems*, Johns Hopkins University, MD, EEUU, Mar 2017.

Personnel Supported by this Grant at Universidad ORT Uruguay

Fernando Paganini.	PI, Professor of Engineering
Andrés Ferragut	Co-PI, Associate Professor

Interactions/Transitions

a) Invited presentation by the PI: "Population dynamics in networks: from queues to PDEs and control, from P2P to deferrable power loads", Institute for Mathematics and its Applications, University of Minnesota. Sept 2015.

b) Presentation of paper [1] by Co-PI at the Allerton Conference, Monticello, Illinois, Oct 2015.

c) Presentation of paper [2] by Co-PI at ACM SIGMETRICS Conference, Juan-Les-Pins, France, Jun 2016.

d) Presentation of poster [3] by Co-PI at Stochastic Networks Conference, San Diego, CA, Jun 2016.

e) Presentation of paper [4] by student I. Rodríguez, ACM SIGCOMM-LANCOMM Workshop, Florianopolis, Brazil, Aug 2016.

f) Invited seminar by the PI, "Controlling populations in networks: a macroscopic view from fluid queues". LIDS-Massachusetts Institute of Technology, Oct. 2016.